

ІННОВАЦІЙНІ ПІДХОДИ ДО СТВОРЕННЯ КОЛЕКТИВНОЇ ПАМ'ЯТІ НА ОСНОВІ ПАРАДИГМИ SEMANTIC WEB

О.В. Новицький

Поняття електронної бібліотеки

Ж Електронна бібліотека це об'єднання через мережу електронних **текстів, документів, зображень, звуків, наукових даних та програмного забезпечення** яке є ядром сьогоденішнього Інтернету, а в майбутньому через організацію доступу до електронних бібліотек буде утворюватися база знань людства.





ПОНЯТТЯ КОЛЕКТИВНОЇ пам'яті

✎ комітет IEEE Technical Committee on Digital Libraries, трактує це поняття як сукупність електронних бібліотек, електронних музеїв, електронних архівів

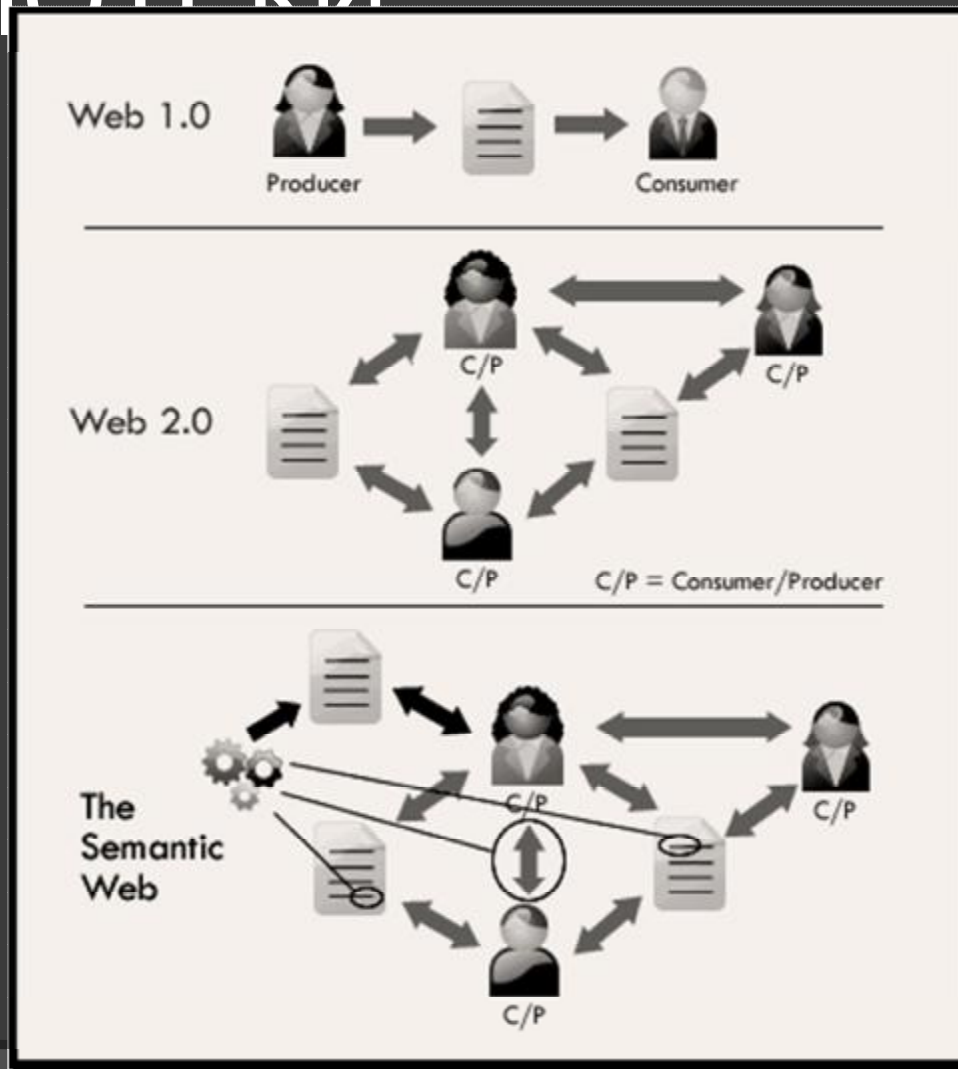
Переваги

- ž дає можливість безпосередньо просліджувати виникнення наукових даних
- ž доступ до інформації будь-яким користувачам будь-де
- ž обмін інформацією дасть можливість аналізувати дані спостережень одночасно багатьом науковим групам навіть якщо вони будуть знаходитися на дуже великих відстанях



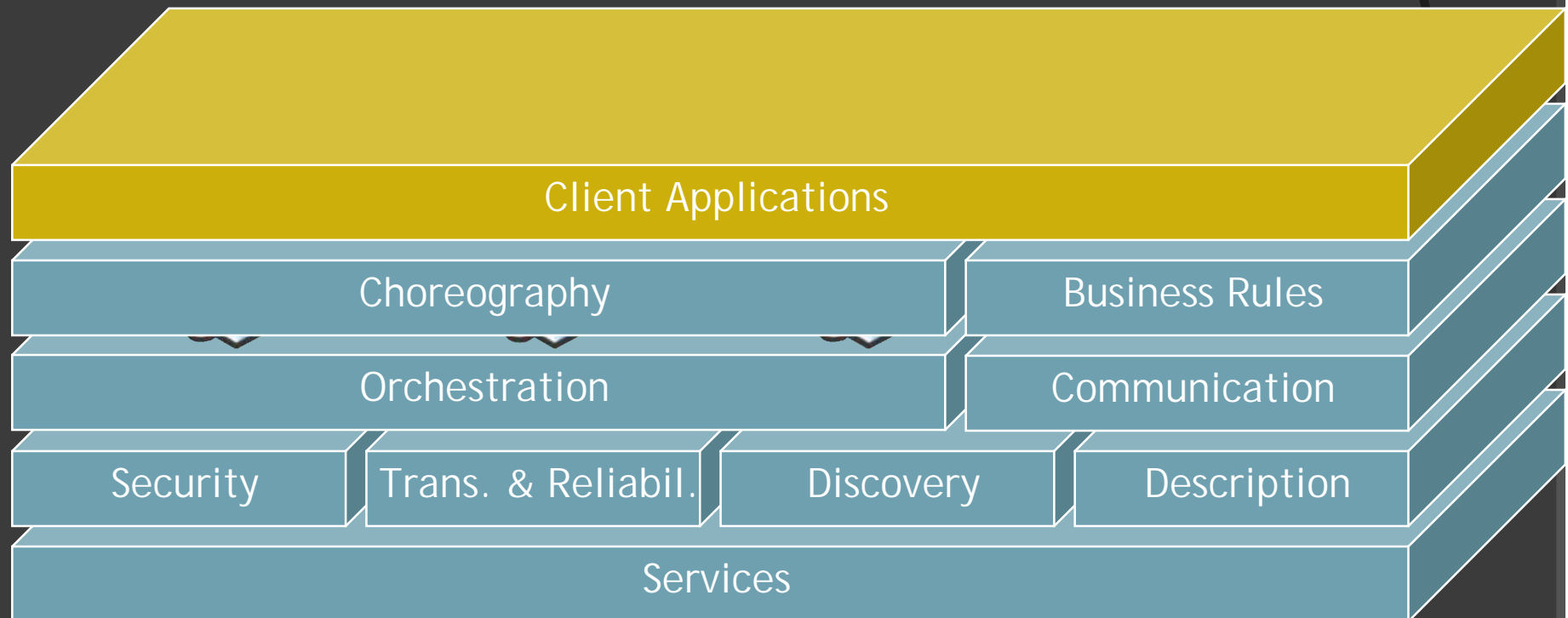
ПАРАДИГМА SEMANTIC WEB В КОНТЕКСТІ ЕЛЕКТРОННОЇ БІБЛІОТЕКИ

1



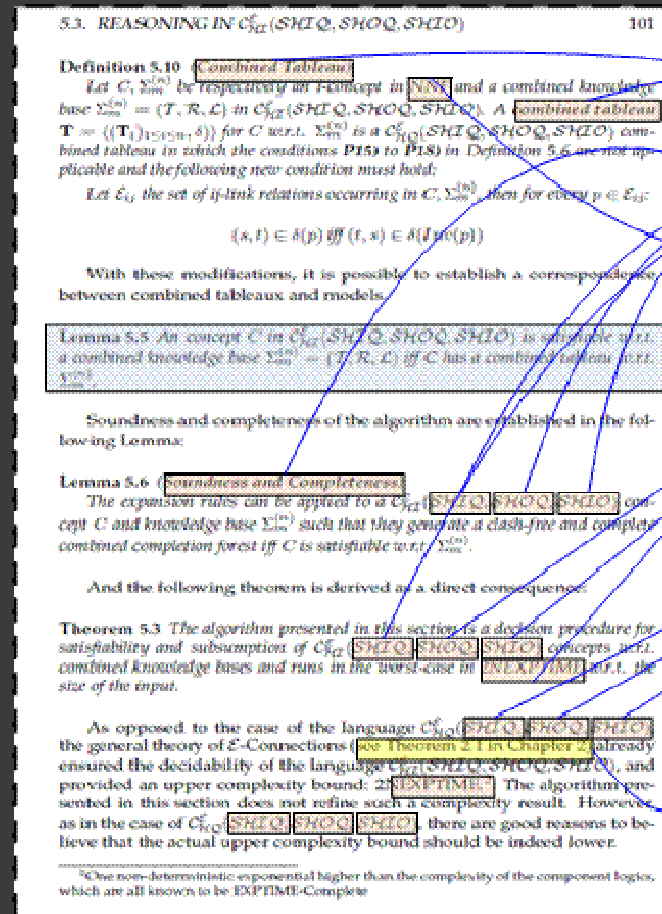
ПАРАДИГМА SEMANTIC WEB В КОНТЕКСТІ ЕЛЕКТРОННОЇ БІБЛІОТЕКИ

2



анотування документа

1



Зовнішні онтології

Інші структурні елементи DO

Метадані

анотація

$$a = \left(\begin{array}{l} h_a \in H(k), \\ A_a \subseteq SN(k) \times LT \times ST(k) \times \\ \times SM(k-1) \times H(k-1) \end{array} \right)$$

LT

множина типів посилань

SN

множину типів анотацій, яка додається до цифрового об'єкту

ST

множина сегментів

SM

множину потоків

H

множина унікальних ідентифікаторів

сервіс

ž Під **сервісом** ми розуміємо деякою послугу мета якої задовольнити запити користувача. Сервіси реалізуються за допомогою **веб-сервісів**. Веб-сервіс це програмне застосування. Основна відмінність між сервісом і веб-сервісом полягає в тому, що той самий сервіс одночасно можна реалізувати **різними** веб-сервісами.



сенсорні сервіси

✂ Сервіси які не впливають інформаційне наповнення ЕБ будемо називати **сенсорні** сервіси.

✂ *поняття сенсорних сервісів для ЕБ ми ввели тому, що ці сервіси аналогічно до сенсорів фільтрують інформаційне середовище, фактично не впливаючи на зміст цього середовища.*



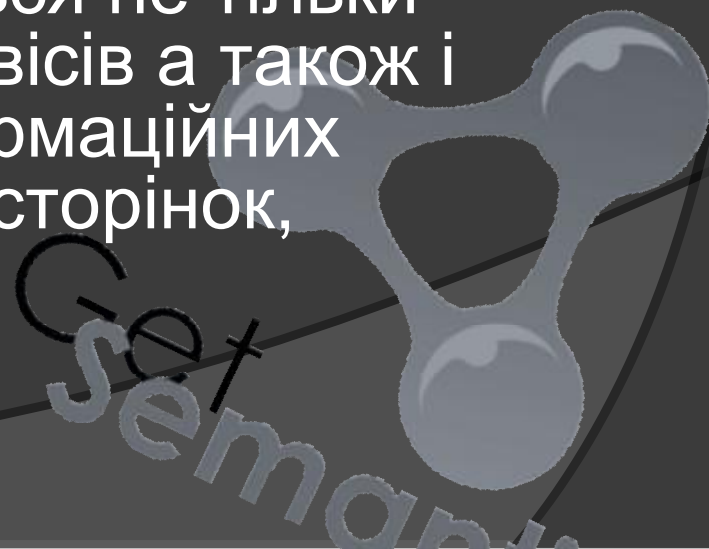
сервіси

- ✂ Сервіс локалізації. Визначає місцезнаходження заданого об'єкту і шлях до нього.
- ✂ Сервіс базового або простого пошуку. Забезпечує пошук по базових наборах метаданих.
- ✂ Сервіс розширеного пошуку. Забезпечує пошук по всіх полях логічним комбінаціям полів метаданих.
- ✂ Сервіс навігації по року видання. Займається побудовою дерева. Будує дерево на основі року видання.
- ✂ Сервіс навігації по предметному покажчику. Будує дерево на основі предметного покажчика.
- ✂ Сервіс навігації по структурних підрозділах. Будує дерево на основі структурних підрозділів.
- ✂ Сервіс навігації по авторах. Будує дерево на основі авторів або творців об'єктів.
- ✂ Всі наступні сервіси не відносяться до категорії сенсорних:
- ✂ Сервіс перетворення форматів. Призначений для перетворення документів одного формату в інший. Цей сервіс є необхідним, оскільки він забезпечить роботу бібліотеки з різними форматами даних.
- ✂ Сервіс індексу-цитовання. Забезпечує визначення рівня індексу цитування документа. Тобто на основі аналізу тексту, будується карта документів і об'єктів, на які посилається вихідний документ.
- ✂ Сервіс побудови графа цитування. Виконує функції побудови графа цитування, де вузлами графа будуть об'єкти або документи, а зв'язки між вузлами є власне посилання. При побудові семантичного графа вузлами можуть виступати, як ресурс або частка ресурсу, а також елементи бази знань, а зв'язки між вузлами виступають як стосунки між об'єктами.
- ✂ Сервіс індексації. Відповідає за індексування ресурсів з урахуванням лексичних форм. Він також допускає вибіркове індексування об'єктів.
- ✂ Сервіс конвертації метаданих. Відповідає за конвертацію одних схем метаданих в інших.
- ✂ Сервіс управління ресурсами. Забезпечить управління авторським правом, політиками і дублікатами об'єктів електронної бібліотеки.
- ✂ Сервіс ідентифікації двійників ресурсів. Сервіс призначений для визначення схожості двох ресурсів. Це досить важливий сервіс, оскільки він виявляє перекриття між будь-якими частками ресурсу, тому дозволяє відшукувати дублікати на будь-якому рівні. Крім того, з використанням технології Semantic Web, чтоне можливим визначати перекриття ресурсів певної наукової області.
- ✂ Сервіс зберігання репозиторія. Займається зберіганням і резервним копіюванням.
- ✂ Сервіс збереження мультимедійної інформації. Відповідає за збереження і завантаження потокової інформації.
- ✂ Сервіс управління об'єктами. Відповідає за уявлення, видання і заміну документів.
- ✂ Сервіс управління доступом. Забезпечує авторизацію користувачів, а також відновлення і встановлення паролів доступу. Також забезпечує багаторівневу систему доступу до інформації.
- ✂ Сервіс управління користувачами. Призначений для управління правами групи користувачів системи.
- ✂ Сервіс збереження результатів пошуку. Призначений для користувачів, які пройшли авторизацію, причому можливо вибіркове збереження результатів.
- ✂ Сервіс інтерфейсу користувача. Відповідає за взаємодію між прикладними сервісами і протоколами. Кожне таке обслуговування може бути централізованим або розподіленим.
- ✂ Сервіс історії даних. Описує всі передумови і весь шлях виникнення і зміни даних.
- ✂ Сервіс зворотного зв'язку. Забезпечує можливість зворотного зв'язку з автором об'єкту, на основі будь-якої доступної контактної інформації, яку вказав автор.
- ✂ Сервіс анотування інформації. Сервіс анотування дозволяє анотувати ресурси та формувати семантичні моделі DO. Ми вважаємо, що з огляду на розвиток технології Semantic Web в основному всі DO будуть мати свої семантичні моделі.
- ✂ Сервіс інтеграції. Виконує роль побудови колективної пам'яті, шляхом об'єднання електронних бібліотек.



Semantic Web

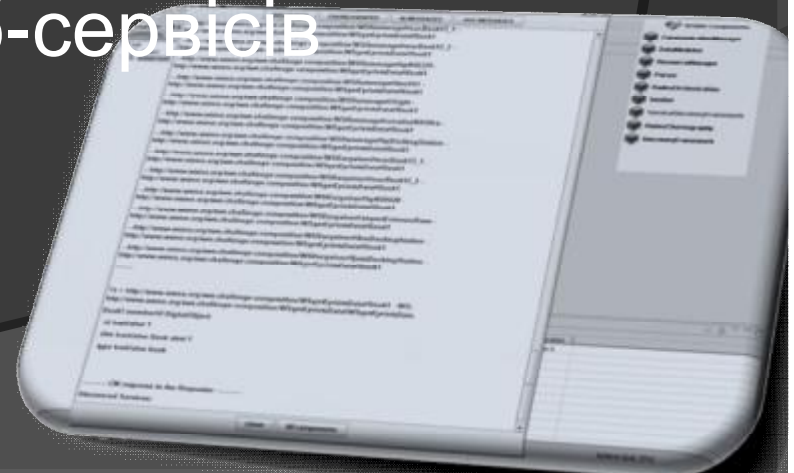
ž Для того, щоб забезпечити динамічне налаштування електронної бібліотеки, необхідно забезпечити механізми, які виконуватимуть цю функцію. Таким механізмом є введення семантики в середовище функціонування веб-сервісів. Цей підхід отримав назву **Semantic Web**, і стосується не тільки семантичного опису сервісів а також і семантичного опису інформаційних ресурсів Інтернету (веб-сторінок, документів тощо.)



WSD та WSC



- ž Проблема пошуку веб-сервісу, який може самотійно виконати запит носить назву дослідження веб-сервісів (Web Service Discovery - **WSD**)
- ž Web Service Composition - **WSC** це проблема композиції веб-сервісів



маштабні задачі WSC

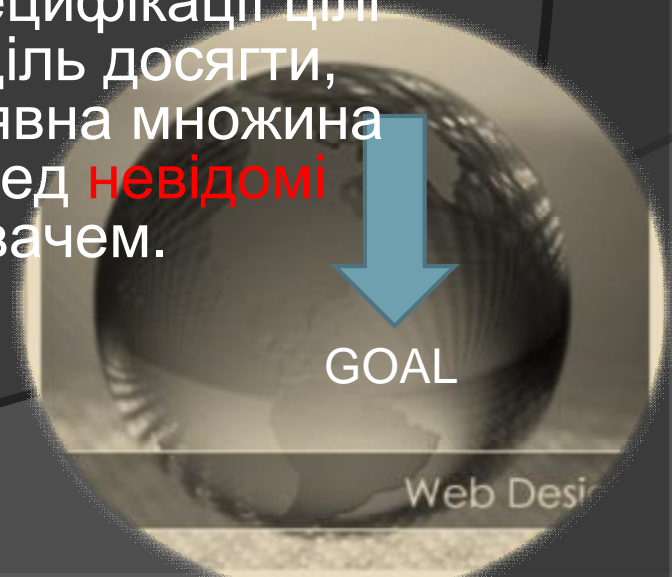
✎ Для велико-маштабних задач WSC алгоритми повного перебору не є ефективними, в даному випадку більш придатними є алгоритми автоматичного пошуку підоптимальних рішень вибірки сервісів на основі знань про предметну область. Введення **семантичного опису** веб-сервісів дасть можливість автоматично вирішувати задачі WSC та WSD.

аналіз факторів які впливають на середовище

- ž в процесі функціонування сервіс-орієнтованої електронної бібліотеки середовище в якому функціонують веб-сервіси постійно змінюється.
- ž **Причини**
- ž проблеми, які породжують розподілені системи часові затримки і ненадійність транспортного протоколу
- ž **людський фактор** – можуть змінитися вимоги користувача, а отже це певним чином впливатиме на результат. Пересічні користувачі цільової аудиторії **не мають чіткого уявлення** про архітектуру бібліотеки, а отже **не можуть** наперед чітко визначитися з цілями які повинні задовольняти сервіси.

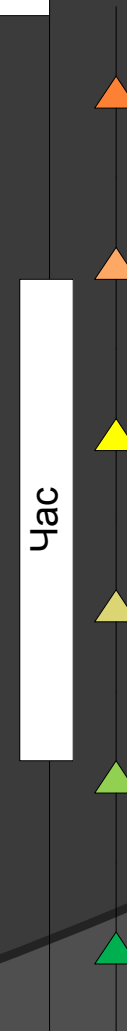
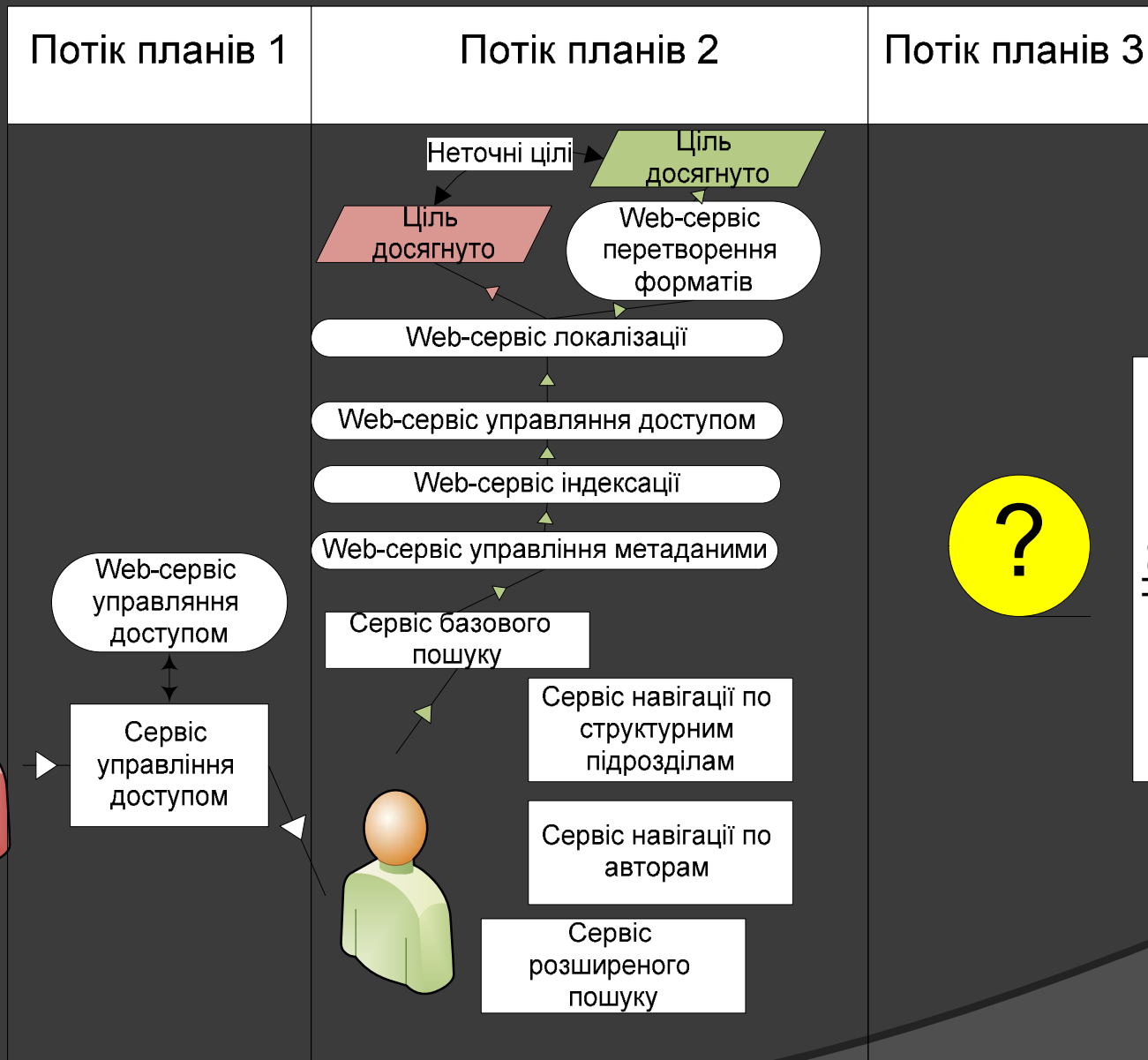
ціле-орієнтована парадигма

- ✎ Ми будемо розглядати композицію на основі ціле-орієнтованої парадигми, тобто виходячи початкових умов та наявної множини сервісів здійснити композицію.
- ✎ Причому, оскільки веб-сервіси розміщені в семантичному середовищі, то вибирати такі плани композиції, які можуть бути корисними для кінцевого користувача. Тобто на відміну від класичної постановки задачі, від специфікації цілі до пошуку сервісів, які зможуть цю ціль досягти, виходити з того припущення, що наявна множина сервісів може досягнути **деякі** наперед **невідомі цілі**, які можуть бути **обрані** користувачем.



wicked задачі

- ž В складній системі цілі, які постають перед композицією можуть змінюватися.
- ž Ми також виходимо з того, що знання про оточуючий світ є **не повними** а отже і цілі є неповними. Для таких конфліктних цілей Horst Rittel і Melvin Webber ввели поняття **вікід** (wicked) **задачі**

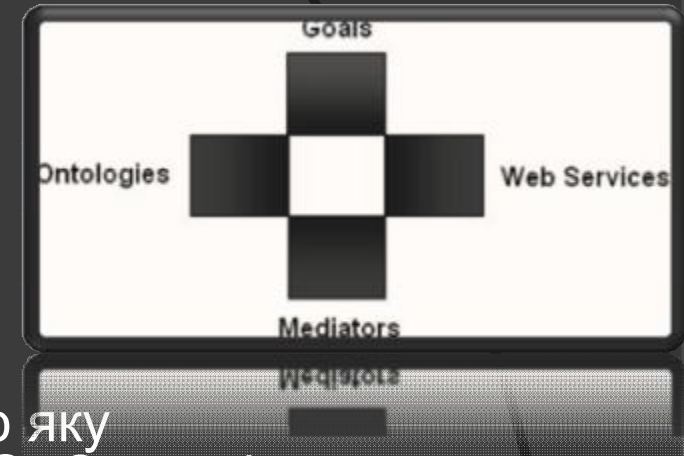


ОСОБЛИВОСТІ ПОБУДОВИ СЕМАНТИЧНОГО СЕРЕДОВИЩА



ž **WSMO** (WSMO - Web Service Modeling Ontology) підхід в якому відображено всі аспекти які пов'язані з семантичними сервісами, які доступні через веб інтерфейс. Кінцевою ціллю цього підходу являється представлення інформації для автоматичного машинного вирішення задач дослідження, вибору, композиції, співставлення, виконання та моніторингу

КОМПОНЕНТИ WSMO



- ž WSMO оперує 4 головними поняттями:
- ž **Онтології** представляють термінологію яку використовують інші компоненти WSMO. Онтологія описує предметну область в якій функціонують веб-сервіси
- ž **Веб-сервіси** представляють обчислювальні об'єкти, таким чином, що в деякій предметній області ці об'єкти являють собою сервіси. Веб-сервіси включає в себе можливості, інтерфейси, внутрішню роботу все це описано в рамках термінології визначеної онтологіями.
- ž **Цілі** представляють побажання користувача, відповідно для задоволення яких можна відшукати веб-сервіс.
- ž **Посередник** – описує ті елементи, які відповідають за вирішення питань функціональної сумісності між іншими елементам WSMO..

WSML

✎ Синтаксис WSMO визначається мовою WSML - Web Service Modeling Language.



Capability та Interface

- ž Специфікація сервісів складається з специфікації можливостей (Capability) та інтерфейсів (Interface) в WSMML.
- ž В WSMML ціль може вимагати декілька інтерфейсів і веб-сервіс може надавати декілька інтерфейсів. Хоча специфікація інтерфейсів не є обов'язковою, з точки зору ціле-орієнтованої парадигми визначення інтерфейсів є **обов'язковим** елементом.

Хореографія

- ž Хореографія інтерфейсу веб-сервісу WSMO, описує поведінку веб-сервісу з точки зору клієнта, такий підхід знаходиться у відповідності до визначення у <http://www.w3.org/TR/ws-gloss/>. Хореографія WSMO керується через механізми станів, і використовує методологію машини Абстрактних станів (Abstract State Machine – ASM)

Особливості формування цілей

- ž В рамках нашої моделі транзитивні правила та умови формують цілі композиції веб-сервісів. Оскільки транзитивні правила визначають можливі переходи між станами, загальна семантика цих переходів визначається в цілях. Таким чином цілі в WSMO **повинні визначати обмеження можливих переходів станів внаслідок виконання транзитивних правил**. Тобто це дасть втілити припущення, що наявна множина сервісів може досягнути деякі наперед невідомі цілі, які можуть бути обрані користувачем. Очевидно, що чим слабше специфікована ціль тим простіше її досягнути.

Наслідки

- ✧ Це дозволить значно покращити функціональність та якість обслуговування кінцевих користувачів електронної бібліотеки, оскільки система сама буде визначати можливі сервіси які будуть надаватися користувачу, а поряд виникає можливість індивідуального обслуговування. В наслідок цього породжується поняття *індивідуальної електронної бібліотеки*, що передбачає не тільки індивідуалізацію контенту але й індивідуалізацію сервісів для роботи з цим контентом.

Дякую за увагу

- ž Більше інформації можна отримати на сайтах:
- ž <http://eprints.zu.edu.ua>
- ž <http://eprints.isofts.kiev.ua>